



The use of a genetic modification in the gene for human G protein $\beta 3$ subunit for the diagnosis of diseases

- 5 The present invention relates to a method for the diagnosis of diseases by genetic analysis, in particular the analysis of genes for subunits of the human guanine nucleotide-binding proteins (G proteins).
- 10 Heterotrimeric guanine nucleotide-binding proteins (G proteins) have an outstanding importance in intracellular signal transduction. They mediate the relaying of extracellular signals after stimulation of hormone receptors and other receptors which undergo a conformational change after receptor activation. This
- 15 leads to activation of G proteins which may subsequently activate or inhibit intracellular effectors (eg. ion channels, enzymes). Heterotrimeric G proteins consist of three subunits, the α , β and γ subunits. To date, several different α subunits, 5 β subunits and about 12 γ subunits have been detected by biochemical and mo-
- 20 lecular biological methods (Birnbaumer, L. and Birnbaumer, M. Signal transduction by G proteins: 1994 edition. *J.Recept.Res.* 15:213-252, 1995; Offermanns, S. and Schultz, G. Complex information processing by the transmembrane signaling system involving G proteins. *Naunyn Schmiedebergs Arch.Pharmacol.* 350:329-338, 1994;
- 25 Nürnberg, B., Gudermann, T., and Schultz, G. Receptors and G proteins as primary components of transmembrane signal transduction. Part 2. G proteins: structure and function. *J.Mol.Med.* 73:123-132, 1995; Neer, E.J. Heterotrimeric G proteins: Organizers of Transmembrane Signals. *Cell* 80:249-257, 1995; Rens-Do-
- 30 miano, S. and Hamm, H.E. Structural and functional relationships of heterotrimeric G-proteins. *FASEB J.* 9:1059-1066, 1995).

- Receptor-mediated activation of certain α subunits can be inhibited by pretreatment with pertussis toxin (PTX). These
- 35 include, in particular, the α isoforms $\alpha i1$, $\alpha i2$ and $\alpha i3$, and various α subunits. G proteins of these types are also referred to as PTX-sensitive G proteins.

- We have found that a genetic modification in the gene for human
- 40 G protein $\beta 3$ subunits is suitable for the diagnosis of diseases. This genetic modification is particularly suitable for establishing the risk of developing a disorder associated with G protein dysregulation.

- 45 The invention furthermore relates to a method for establishing a relative risk of developing disorders associated with G protein dysregulation for a subject, which comprises comparing the gene

sequence for human G protein $\beta 3$ subunit of the subject with the gene sequence SEQ ID NO:1, and, in the event that a thymine (T) is present at position 825, assigning the subject an increased risk of disease.

5

The genetic modification which has been found is located in the gene for human G protein $\beta 3$ subunit. This gene has been described by Levine et al. (Proc. Natl. Acad. Sci USA, 87, (1990) 2329-2333). The coding region has an Ser codon (TCC) at position

- 10 275, while subjects with an increased risk of a disease associated with G protein dysregulation have the codon TCT, which likewise codes for Ser, at this position. The genetic modification is a base substitution at position 825 in which a cytosine (C) is replaced by thymine (T). However, this base
- 15 exchange is "silent" at the amino-acid level, ie. it does not lead to incorporation of a different amino acid at this position. The sequence found in subjects with an increased risk of disease is depicted in SEQ ID NO:1 in the sequence listing.

- 20 The genetic modification which has been found usually occurs in heterozygous form.

Disorders associated with G protein dysregulation are defined as diseases in which the G protein is involved in signal

- 25 transduction and does not carry out its function in a physiological manner.

The dysregulation may have a number of causes, for example a modification in the structural gene or modified gene expression.

30

The disorders include cardiovascular diseases, metabolic disturbances and immunological diseases.

Cardiovascular diseases which may be mentioned are:

35

Hypertension, pregnancy hypertension (gestosis, hypertension in pregnancy), coronary heart disease, localized and/or generalized atherosclerosis, stenoses of blood vessels, restenosis after revascularizing procedures (eg. PTCA with and without stent

40 implantation), tendency to stroke or thrombosis and increased platelet aggregation.

Metabolic disturbances which may be mentioned are:

- 45 Metabolic syndrome, insulin resistance and hyperinsulinemia, type II diabetes mellitus, diabetic complications (eg. nephropathy, neuropathy, retinopathy, etc.) disturbances of lipid

metabolism, disturbances of central chemoreception (CO₂ tolerance, acidosis tolerance, sudden infant death (SIDS)).

Immunological diseases which may be mentioned are:

5

Impaired strength of the body's immune response (formation of immunoglobulins, aggressiveness of T cells and NK cells), impaired general tendency to proliferation, including wound-healing capacity, tendency to develop tumors and

10 proliferation including metastasizing potential of malignantly transformed cells, duration of the latency period after HIV infection until the disease becomes clinically evident, Kaposi sarcoma, tendency to cirrhosis of the liver, transplant tolerance and transplant rejection.

15

The use of the genetic mutation according to the invention is particularly suitable for establishing the risk of developing hypertension.

20 The invention furthermore relates to the production of transgenic animals harboring the genetic mutation described above.

Transgenic animals of this type are of great importance in particular as animal models for the investigation and therapy of the disorders described above. The methods for generating

25 transgenic animals are generally known to the skilled worker.

For the method according to the invention for establishing the relative risk of developing a disease, body material containing the subject's genetic information is taken from a subject. This

30 is achieved as a rule by taking blood and isolating the nucleic acid therefrom.

The structure of the gene for the G protein $\beta 3$ subunit is established from the subject's isolated nucleic acid and is

35 compared with the sequence indicated in SEQ ID NO:1.

The structure of the gene can be established by sequencing of the nucleic acid. This can take place either directly from the genomic DNA or after amplification of the nucleic acid, for

40 example by the PCR technique.

The structure of the gene can take place at the genomic level or else at the mRNA or cDNA level.

45 It is preferably established by sequencing after PCR amplification of the cDNA. The primers suitable for the PCR can easily be inferred by the skilled worker from the sequences

depicted in SEQ ID NO:1. The procedure for this is advantageously such that in each case a primer binding a strand and complementary strand in front of and behind the relevant base position 825 is chosen.

5

However, other methods can also be used for comparison of the genes, for example selective hybridization or appropriate mapping with restriction enzymes. The C→T base exchange at the position 825 described above leads to loss of a cleavage site for the
10 restriction enzyme Dsa I, which is likewise used to detect this genetic polymorphism.

If the subject has a thymine (T) at position 825, he is to be assigned a greater risk of disease than a subject with a cytosine
15 (C) at this position.

The invention is illustrated further in the following examples.

Example 1

20

Detection of the genetic modification in hypertensives by sequencing

An enhanced susceptibility to activation of PTX-sensitive G
25 proteins was detected in preliminary investigations on patients with essential hypertension. This detection was possible in immortalized cells from patients having as phenotypical marker an enhanced activity of the Na/H exchanger. The enhanced susceptibility to activation of PTX-sensitive G proteins has
30 important consequences for cellular function. These include enhanced formation of intracellular second messenger molecules (eg. inositol 1,4,5-trisphosphate), enhanced release of intracellular Ca^{2+} ions, increased formation of immunoglobulins and an increased rate of cell growth. Since these changes can be
35 detected in immortalized cells and after a long duration of cell culturing, it may be assumed that this modification is genetically fixed (Rosskopf, D., Frömter, E., and Siffert, W. Hypertensive sodium-proton exchanger phenotype persists in immortalized lymphoblasts from essential hypertensive patients—a cell
40 culture model for human hypertension. *J.Clin.Invest.* 92:2553-2559, 1993; Rosskopf, D., Hartung, K., Hense, J., and Siffert, W. Enhanced immunoglobulin formation of immortalized B cells from hypertensive patients. *Hypertension* 26:432-435, 1995; Rosskopf, D., Schröder, K.-J., and Siffert, W. Role of sodium-hydrogen exchange in the proliferation of immortalised lymphoblasts
45 from patients with essential hypertension and normotensive subjects. *Cardiovasc.Res.* 29:254-259, 1995; Siffert, W., Rosskopf,

D., Moritz, A., Wieland, T., Kaldenberg-Stasch, S., Kettler, N., Hartung, K., Beckmann, S., and Jakobs, K.H. Enhanced G protein activation in immortalized lymphoblasts from patients with essential hypertension. *J.Clin.Invest.* 96:759-766, 1995).

5

RNA was prepared by standard methods from immortalized cell lines from hypertensives and was transcribed into cDNA using reverse transcriptase. Using the polymerase chain reaction (PCR), the cDNA coding for the G protein $\beta 3$ subunit was amplified and

10 sequenced. The following oligonucleotide primers were employed for the PCR:

5'-TGG GGG AGA TGG AGC AAC TG and

5'-CTG CTG AGT GTG TTC ACT GCC.

15

Compared with the sequence published by Levine et al. (Levine, M.A., Smallwood, P.M., Moen, P.T., Jr., Helman, L.J., and Ahn, T.G. Molecular cloning of $\beta 3$ subunit, a third form of the G protein β -subunit polypeptide. *Proc. Natl. Acad. Sci. USA*

20 87(6):2329-2333, 1990), the following difference was found in the cDNA from hypertensives' cells: nucleotide 825 cytosine (C) in the region of the coding sequence is replaced by a thymine (T) (nucleotide 1 corresponds to base A in the ATG start codon). This base exchange leads to a silent polymorphism, ie. the amino

25 acid encoded by the corresponding base triplet (serine) is not altered by comparison with the original sequence. The DNA sequence found is described in SEQ ID NO:1.

Example 2

30

Detection of the genetic modification in hypertensives by restriction enzyme analysis

The figure depicts a comparison of genes from normotensives and

35 hypertensives by restriction enzyme analysis. In this, the cDNA coding for $\beta 3$ from cells from normotensives (NT) and hypertensives (HT), which had been amplified by PCR, was subjected to a restriction enzyme analysis using the enzyme Dsa I. The reaction products were fractionated in an agarose gel, which is depicted

40 in the figure.

The complete restriction of $\beta 3$ cDNA from normotensive cells after digestion with Dsa I is clearly evident from the figure. The cDNA from hypertensives' cells is only partly cut by Dsa I. Apart from

45 the cleavage products to be expected there is also uncleaved PCR product. Reference fragments (markers) are loaded on the left and

right for comparison of sizes. Four of the five DNA sequences from hypertensives depicted here show the base exchange described above and are heterozygous for this modification.

5

10

15

20

25

30

35

40

45

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: BASF Aktiengesellschaft
- (B) STREET: Carl-Bosch-Strasse 38
- (C) CITY: Ludwigshafen
- (E) COUNTRY: Federal Republic of Germany
- (F) POSTAL CODE: D-67056
- (G) TELEPHONE: 0621/6048526
- (H) TELEFAX: 0621/6043123
- (I) TELEX: 1762175170

(ii) TITLE OF APPLICATION: Method for diagnosing disorders by analysis of genes

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1517 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA for mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURES:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1024

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | GGG | GAG | ATG | GAG | CAA | CTG | CGT | CAG | GAA | GCG | GAG | CAG | CTC | AAG | AAG | 48 |
| Met | Gly | Glu | Met | Glu | Gln | Leu | Arg | Gln | Glu | Ala | Glu | Gln | Leu | Lys | Lys | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | |
| CAG | ATT | GCA | GAT | GCC | AGG | AAA | GCC | TGT | GCT | GAC | GTT | ACT | CTG | GCA | GAG | 96 |
| Gln | Ile | Ala | Asp | Ala | Arg | Lys | Ala | Cys | Ala | Asp | Val | Thr | Leu | Ala | Glu | |
| | | | 20 | | | | 25 | | | | | 30 | | | | |
| CTG | GTG | TCT | GGC | CTA | GAG | GTG | GTG | GGA | CGA | GTC | CAG | ATG | CGG | ACG | CGG | 144 |
| Leu | Val | Ser | Gly | Leu | Glu | Val | Val | Gly | Arg | Val | Gln | Met | Arg | Thr | Arg | |
| | | | 35 | | | | 40 | | | | | 45 | | | | |
| CGG | ACG | TTA | AGG | GGA | CAC | CTG | GCC | AAG | ATT | TAC | GCC | ATG | CAC | TGG | GCC | 192 |
| Arg | Thr | Leu | Arg | Gly | His | Leu | Ala | Lys | Ile | Tyr | Ala | Met | His | Trp | Ala | |
| | | | 50 | | | 55 | | | | 60 | | | | | | |
| ACT | GAT | TCT | AAG | CTG | CTG | GTA | AGT | GCC | TCG | CAA | GAT | GGG | AAG | CTG | ATC | 240 |
| Thr | Asp | Ser | Lys | Leu | Leu | Val | Ser | Ala | Ser | Gln | Asp | Gly | Lys | Leu | Ile | |
| | | | 65 | | 70 | | | | 75 | | | | 80 | | | |
| GTG | TGG | GAC | AGC | TAC | ACC | ACC | AAC | AAG | GTG | CAC | GCC | ATC | CCA | CTG | CGC | 288 |
| Val | Trp | Asp | Ser | Tyr | Thr | Thr | Asn | Lys | Val | His | Ala | Ile | Pro | Leu | Arg | |
| | | | | 85 | | | 90 | | | | | 95 | | | | |
| TCC | TCC | TGG | GTC | ATG | ACC | TGT | GCC | TAT | GCC | CCA | TCA | GGG | AAC | TTT | GTG | 336 |
| Ser | Ser | Trp | Val | Met | Thr | Cys | Ala | Tyr | Ala | Pro | Ser | Gly | Asn | Phe | Val | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------|-------------|-------------|------------|-------------|------------|------------|------------|------------|------------|------|-----|-----|-----|-----|-----|------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| 100 | | | | | | | | | | | | | | | | 105 | | | | | | | | | | | | | | | | 110 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| GCA | TGT | GGG | GGG | CTG | GAC | AAC | ATG | TGT | TCC | ATC | TAC | AAC | CTC | AAA | TCC | 384 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ala | Cys | Gly | Gly | Leu | Asp | Asn | Met | Cys | Ser | Ile | Tyr | Asn | Leu | Lys | Ser | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 115 | | | | | | | | | | | | | | | | 120 | | | | | | | | | | | | | | | | 125 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CGT | GAG | GGC | AAT | GTC | AAG | GTC | AGC | CGG | GAG | CTT | TCT | GCT | CAC | ACA | GGT | 432 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Arg | Glu | Gly | Asn | Val | Lys | Val | Ser | Arg | Glu | Leu | Ser | Ala | His | Thr | Gly | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 130 | | | | | | | | | | | | | | | | 135 | | | | | | | | | | | | | | | | 140 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TAT | CTC | TCC | TGC | TGC | CGC | TTC | CTG | GAT | GAC | AAC | AAT | ATT | GTG | ACC | AGC | 480 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Tyr | Leu | Ser | Cys | Cys | Arg | Phe | Leu | Asp | Asp | Asn | Asn | Ile | Val | Thr | Ser | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 145 | | | | | | | | | | | | | | | | 150 | | | | | | | | | | | | | | | | 155 | | | | | | | | | | | | | | | | 160 | | | | | | | | | | | | | | | | |
| TCG | GGG | GAC | ACC | ACG | TGT | GCC | TTG | TGG | GAC | ATT | GAG | ACT | GGG | CAG | CAG | 528 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ser | Gly | Asp | Thr | Thr | Cys | Ala | Leu | Trp | Asp | Ile | Glu | Thr | Gly | Gln | Gln | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 165 | | | | | | | | | | | | | | | | 170 | | | | | | | | | | | | | | | | 175 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AAG | ACT | GTA | TTT | GTG | GGA | CAC | ACG | GGT | GAC | TGC | ATG | AGC | CTG | GCT | GTG | 576 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Lys | Thr | Val | Phe | Val | Gly | His | Thr | Gly | Asp | Cys | Met | Ser | Leu | Ala | Val | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 180 | | | | | | | | | | | | | | | | 185 | | | | | | | | | | | | | | | | 190 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TCT | CCT | GAC | TTC | AAT | CTC | TTC | ATT | TCG | GGG | GCC | TGT | GAT | GCC | AGT | GCC | 624 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ser | Pro | Asp | Phe | Asn | Leu | Phe | Ile | Ser | Gly | Ala | Cys | Asp | Ala | Ser | Ala | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 195 | | | | | | | | | | | | | | | | 200 | | | | | | | | | | | | | | | | 205 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AAG | CTC | TGG | GAT | GTG | CGA | GAG | GGG | ACC | TGC | CGT | CAG | ACT | TTC | ACT | GGC | 672 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Lys | Leu | Trp | Asp | Val | Arg | Glu | Gly | Thr | Cys | Arg | Gln | Thr | Phe | Thr | Gly | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 210 | | | | | | | | | | | | | | | | 215 | | | | | | | | | | | | | | | | 220 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CAC | GAG | TCG | GAC | ATC | AAC | GCC | ATC | TGT | TTC | TTC | CCC | AAT | GGA | GAG | GCC | 720 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| His | Glu | Ser | Asp | Ile | Asn | Ala | Ile | Cys | Phe | Phe | Pro | Asn | Gly | Glu | Ala | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 225 | | | | | | | | | | | | | | | | 230 | | | | | | | | | | | | | | | | 235 | | | | | | | | | | | | | | | | 240 | | | | | | | | | | | | | | | | |
| ATC | TGC | ACG | GGC | TCG | GAT | GAC | GCT | TCC | TGC | CGC | TTG | TTT | GAC | CTG | CGG | 768 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ile | Cys | Thr | Gly | Ser | Asp | Asp | Ala | Ser | Cys | Arg | Leu | Phe | Asp | Leu | Arg | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 245 | | | | | | | | | | | | | | | | 250 | | | | | | | | | | | | | | | | 255 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| GCA | GAC | CAG | GAG | CTG | ATC | TGC | TTC | TCC | CAC | GAG | AGC | ATC | ATC | TGC | GGC | 816 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ala | Asp | Gln | Glu | Leu | Ile | Cys | Phe | Ser | His | Glu | Ser | Ile | Ile | Cys | Gly | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 260 | | | | | | | | | | | | | | | | 265 | | | | | | | | | | | | | | | | 270 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ATC | ACG | TCT | GTG | GCC | TTC | TCC | CTC | AGT | GGC | CGC | CTA | CTA | TTC | GCT | GGC | 864 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Ile | Thr | Ser | Val | Ala | Phe | Ser | Leu | Ser | Gly | Arg | Leu | Leu | Phe | Ala | Gly | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 275 | | | | | | | | | | | | | | | | 280 | | | | | | | | | | | | | | | | 285 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TAC | GAC | GAC | TTC | AAC | TGC | AAT | GTC | TGG | GAC | TCC | ATG | AAG | TCT | GAG | CGT | 912 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Tyr | Asp | Asp | Phe | Asn | Cys | Asn | Val | Trp | Asp | Ser | Met | Lys | Ser | Glu | Arg | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 290 | | | | | | | | | | | | | | | | 295 | | | | | | | | | | | | | | | | 300 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| GTG | GGC | ATC | CTC | TCT | GGC | CAC | GAT | AAC | AGG | GTG | AGC | TGC | CTG | GGA | GTC | 960 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Val | Gly | Ile | Leu | Ser | Gly | His | Asp | Asn | Arg | Val | Ser | Cys | Leu | Gly | Val | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 305 | | | | | | | | | | | | | | | | 310 | | | | | | | | | | | | | | | | 315 | | | | | | | | | | | | | | | | 320 | | | | | | | | | | | | | | | | |
| ACA | GCT | GAC | GGG | ATG | GCT | GTG | GCC | ACA | GGT | TCC | TGG | GAC | AGC | TTC | CTC | 1008 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Thr | Ala | Asp | Gly | Met | Ala | Val | Ala | Thr | Gly | Ser | Trp | Asp | Ser | Phe | Leu | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 325 | | | | | | | | | | | | | | | | 330 | | | | | | | | | | | | | | | | 335 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AAA | ATC | TGG | AAC | TGA | G | GAGGCTGGAG | AAAGGGAAGT | GGAAGGCAGT | GAACACACTC | 1064 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Lys | Ile | Trp | Asn | * | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 340 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AGCAGCCCCC | TGCCCCGACCC | CATCTCATTC | AGGTGTTCTC | TTCTATATTC | CGGGTGCCAT | 1124 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TCCCACATAAG | CTTTCTCCTT | TCAGGGGCAGT | GGGGAGCATG | GGACTGTGCC | TTTGGGAGGC | 1184 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AGCATCAGGG | ACACAGGGGC | AAAGAAGTGC | CCCATCTCCT | CCCATGGCCT | TCCCTCCCCA | 1244 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CAGTCCCTCAC | AGCCTCTCCC | TTAATGAGCA | AGGACAACCT | GCCCCCTCCCC | AGCCCTTTGC | 1304 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AGGCCACAGCA | GACTTGAGTC | TGAGGCCCCA | GGCCCTAGGA | TTCTCTCCCC | AGAGCCACTA | 1364 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CCTTTGTCCA | GGCCTGGGTG | GTATAGGGCG | TTTGGCCCTG | TGACTATGGC | TCTGGCACCA | 1424 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CTAGGGTCTCT | GGCCCTCTTC | TTATTCATGC | TTTCTCCTTT | TTCTACCTTT | TTTTCTCTCC | 1484 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TAAGACACCT | GCAATAAAGT | GTAGCACCTT | GGT | | | 1517 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 341 amino acids

(B) TYPE: Amino acid

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Gly Glu Met Glu Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Lys
 1           5           10           15
Gln Ile Ala Asp Ala Arg Lys Ala Cys Ala Asp Val Thr Leu Ala Glu
 20           25           30
Leu Val Ser Gly Leu Glu Val Val Gly Arg Val Gln Met Arg Thr Arg
 35           40           45
Arg Thr Leu Arg Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Ala
 50           55           60
Thr Asp Ser Lys Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile
 65           70           75           80
Val Trp Asp Ser Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg
 85           90           95
Ser Ser Trp Val Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Phe Val
100           105           110
Ala Cys Gly Gly Leu Asp Asn Met Cys Ser Ile Tyr Asn Leu Lys Ser
115           120           125
Arg Glu Gly Asn Val Lys Val Ser Arg Glu Leu Ser Ala His Thr Gly
130           135           140
Tyr Leu Ser Cys Cys Arg Phe Leu Asp Asp Asn Asn Ile Val Thr Ser
145           150           155           160
Ser Gly Asp Thr Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln
165           170           175
Lys Thr Val Phe Val Gly His Thr Gly Asp Cys Met Ser Leu Ala Val
180           185           190
Ser Pro Asp Phe Asn Leu Phe Ile Ser Gly Ala Cys Asp Ala Ser Ala
195           200           205
Lys Leu Trp Asp Val Arg Glu Gly Thr Cys Arg Gln Thr Phe Thr Gly
210           215           220
His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Glu Ala
225           230           235           240
Ile Cys Thr Gly Ser Asp Asp Ala Ser Cys Arg Leu Phe Asp Leu Arg
245           250           255
Ala Asp Gln Glu Leu Ile Cys Phe Ser His Glu Ser Ile Ile Cys Gly
260           265           270
Ile Thr Ser Val Ala Phe Ser Leu Ser Gly Arg Leu Leu Phe Ala Gly
275           280           285
Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ser Met Lys Ser Glu Arg
290           295           300
Val Gly Ile Leu Ser Gly His Asp Asn Arg Val Ser Cys Leu Gly Val
305           310           315           320
Thr Ala Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu
325           330           335
Lys Ile Trp Asn *
340

```